i2b2 AIRWAYS DISEASE DBP: PATIENTS, PHENOTYPES, MODELS, AND DATA VIEWER

Lazarus, Ross[1]; Murphy, Shawn[2]; Zeng, Qing[3]; Sordo, Margarita[3]; Wei, Lee-Jen[4]; Fuhlbrigge, Anne[1]
Signorovitch, James[4]; Churchill, Suzanne[5]; Weiss, Scott[1]; Kohane, Isaac[6]

[1]Channing Laboratory, Brigham and Womens Hospital and Harvard Medical School; [2]Department of Neurology, Massachusetts General Hospital; [3]Decisions Systems Group, Harvard Medical School; [4]Department of Biostatistics, Harvard School of Public Health; [5]i2b2 NCBC; [6]Informatics Program, Children's Hospital, Boston, MA

Keywords: phenotypes, nlp, visualization tools, genetics, pharmacogenetics

Translating basic research into clinical practice in asthma, a common respiratory illness, is the focus of the airways disease DBP for the i2b2 NCBC. Our overall goal is to elucidate genetic variants associated with individual patient responses to common asthma medications, so we are assembling validated, detailed patient characteristics and DNA samples for genotyping.

In collaboration with the Tools core, we have developed a data warehouse (DW), populated from the existing Partners patient data repository. This DW is a functioning prototype for the patient clinical data component of the i2b2 clinical research chart and contains extensive clinical data on more than 97,000 asthmatics. Some important patient characteristics were not available in coded form from the DW. Collaborating NLP experts from the Scientific Core have provided smoking history, medication use and additional comorbidity measures derived from progress notes and discharge summary texts. They have developed a novel tool which allows generalization of their work, HITEx, which will soon be made freely available as part of the i2b2 toolkit.

We are beginning to directly measure relevant patient characteristics, and store DNA samples for future genotyping, from a substantial sample of Partners asthma patients. The cohort is anticipated to be between 1500 and 2000 subjects, recruited through collaboration with physicians located at Partners asthma outpatient specialist services. We are measuring lung function and bronchodilator responsiveness (BDR), and obtaining a past medical history, smoking history, and medication use history as well as a sample of blood for extraction and storage of genetic material. In addition to serving as phenotypes in our genetic analyses, direct measurement of patient characteristics will be used to validate the DW data. Most of these patients will stay in the system and clinical information about them will continue to accrue, providing a rich source of patient phenotype data, which will allow future genetic and pharmacogenetic studies based on the stored genetic material. For example, genetic variants contributing to variation in BDR may be useful for predicting responsiveness, allowing more appropriately targeted medication.

Some asthmatics experience serious and sudden worsening of their illness. These exacerbations can become life threatening and may be preventable. In collaboration with statisticians from the Scientific core, we have developed models predicting risk of acute exacerbation among childhood asthmatics, using detailed and high-quality measurements obtained as part of a clinical trial (CAMP). Much of the information about exacerbation risk can be obtained from demography and past history. BD responsiveness gave additional useful information and specialised measurements gave only small additional incremental information gain. We will test these models in the Partners clinical data repository, and among our cohort to evaluate their generalisability.

Finally, in developing statistical models and evaluating the NLP phenotypes, it quickly became clear that a novel visualization tool was required. A prototype data viewer for the i2b2 clinical research chart developed by the Tools core, is now in production use for data validation and for exploring temporal relationships between events for groups of asthmatic patients. Supported by grant U54LM008748.